Copyright © 2002 by the Genetics Society of America

Crossover Interference in Arabidopsis

G. P. Copenhaver,*[,1] E. A. Housworth[†,1] and F. W. Stahl[‡,1,2]

**Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, [†]Mathematics Department, University of Oregon, Eugene, Oregon 97403 and [‡]Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403-1229*

Manuscript received October 19, 2001
Accepted for publication January 24, 2002

ABSTRACT

The crossover distribution in meiotic tetrads of *Arabidopsis thaliana* differs from those previously described for Drosophila and Neurospora. Whereas a chi-square distribution with an even number of degrees of freedom provides a good fit for the latter organisms, the fit for Arabidopsis was substantially improved by assuming an additional set of crossovers sprinkled, at random, among those distributed as per chi square. This result is compatible with the view that Arabidopsis has two pathways for meiotic crossing over, only one of which is subject to interference. The results further suggest that Arabidopsis meiosis has >10 times as many double-strand breaks as crossovers.

COBBS (1978) and STAM (1979) proposed tidy mathematical models for crossover (chiasma) interference in meiosis. Their equivalent models envisioned Poisson-distributed crossover "attempts" among acts of meiosis for a specified bivalent. Successful attempts (resulting in crossing over) were separated by a fixed number of "failures," which gave no crossing over. The models are sometimes referred to as "chi-square" models because the resulting probability distribution for interexchange distances is a chi-square distribution with an even number of degrees of freedom (LANGE *et al.* 1997).

FOSS *et al.* (1993), inspired by MORTIMER and FOGEL (1974), offered a "counting model" for crossover interference, assuming no chromatid interference, which (unwittingly) expanded on that of Cobbs and Stam by detailing the nature of the "failures" that are "counted" between crossovers. The failures were assumed to be double-strand breaks that were repaired without crossing over ("noncrossovers"; SZOSTAK *et al.* 1983). Since double-strand-break repair can result in gene conversion whether or not it is accompanied by crossing over, it was reasonable to estimate the number of obligate failures between successes from the fraction of gene conversions that are unaccompanied by crossing over of markers flanking the conversion. For Drosophila, the only estimate of this fraction set the counting number at four (HILLIKER and CHOVNICK 1981; HILLIKER *et al.* 1991), while data for Neurospora gave a value of two (PERKINS *et al.* 1993). A mark of the model's success was that the values four and two for those two organisms, respectively, generated optimal expressions for multiple crossover data analyzed in a variety of ways (FOSS *et al.* 1993; LANDE and STAHL 1993; MCPEEK and SPEED 1995; ZHAO *et al.* 1995). The model received further encouragement from evidence in the literature, albeit weak, of negative interference between crossovers and noncrossovers (see FOSS *et al.* 1993).

A prediction of the counting model, related to the negative interference between crossovers and noncrossovers, is that the interval between a pair of close exchanges should be especially enriched for noncrossovers, with some of them manifested as conversions when markers are present to detect them. Experiments in budding yeast by FOSS and STAHL (1995) failed to support that prediction. Since that time, two developments in yeast genetics imply that the test was doomed: (1) In addition to crossovers that are subject to interference, yeast may have additional crossovers, not subject to interference, that derive from recombinational events required for chromosome pairing. As proposed by ZALEVSKY *et al.* (1999), these crossovers are the ones remaining in *msh4* and *zip1* mutants, which have reduced levels of crossing over and of interference. The existence of such "contaminating" crossovers would confound the quantitative predictions of the counting model. (2) More importantly, the genetic markers used to detect and enumerate the expected noncrossovers may have altered the events such that many were not detected as conversions (BORTS and HABER 1989; BORTS *et al.* 1990; CHEN and JINKS-ROBERTSON 1999). For instance, the chromosomes may have been driven by the presence of the markers to repair some of their breaks using sister chromatids as template. Alternatively, the markers may have prevented breaks.

Why do some organisms (*e.g.*, Drosophila and Neurospora) appear to have crossing over that is subject to simple rules of interference, while interference in yeast appears to be complicated by noninterfering exchanges

[1] The authors contributed equally to this work.

[2] *Corresponding author:* Institute of Molecular Biology, 1229 University of Oregon, 1370 Franklin Blvd., Eugene, OR 97403-1229. E-mail: fstahl@molbio.uoregon.edu

involved in pairing? A provisional answer to this query (Zalevsky *et al.* 1999) lies in the different ways that yeast, on the one hand, and flies, on the other, secure chromosome pairing during prophase I of meiosis. The chromosomes of Drosophila (Hawley 1980; reviewed in McKee *et al.* 2000) and of *Caenorhabditis elegans* (Villeneuve 1994; Albertson *et al.* 1997), which also has robust interference, have *cis*-acting "pairing centers," sequences that ensure homolog pairing. Yeast (for reviews, see Kleckner 1996; Roeder 1997), on the other hand, like mouse (Romanienko and Camerini-Otero 2000) and Coprinus (Celerin *et al.* 2000), relies on recombinational interactions to secure stable pairing. In yeast, these interactions require the meiosis-specific strand-invasion protein Dmc1p (Bishop *et al.* 1992), which appears to be lacking in both Drosophila and *C. elegans*, organisms that apparently employ only the generalized strand-invasion protein Rad51p (Shinohara *et al.* 1992) to effect meiotic double-strand-break repair. This correlation suggests the "rule" that organisms possessing Dmc1p use a noninterference crossover pathway to ensure chromosome pairing, while organisms that lack Dmc1p manifest robust interference of an uncomplicated sort because their chromosomes are paired by a nonrecombinational route.

Arabidosis requires the early recombination function Spo11 to achieve synapsis (Grelon *et al.* 2001), putting it in the camp with yeast, mouse, and Coprinus. That observation plus the demonstrated presence of a *DMC1* homolog in Arabidopsis (Klimyuk and Jones 1997; Doutriaux *et al.* 1998; Couteau *et al.* 1999) requires this green plant, if it is to follow the rule, to have both an interference and a noninterference pathway for meiotic crossing over. This possibility can be tested by scoring the segregation of abundant PCR-based molecular markers in the meiotic tetrads produced in *quartet* mutants of Arabidopsis (Preuss *et al.* 1994; Copenhaver *et al.* 1998).

Our analysis is based on the simple assumption that the disposition of exchange points in the interference pathway is governed by the counting model and that additional exchanges, arising in the pairing pathway, are (pre)sprinkled randomly (*i.e.*, without interference) on this background. The adequacy of our model is supported by control analyses of Neurospora and Drosophila data.

RESULTS

The markers (chromosome 1, nga59, nga63, g2395, m235, SO392, 7G6, T27K12, nga280, ETR, TAG, AthATPASE, nga692; chromsome 2, nga1145, mi310, THY1B, nga1126, nga361, nga168; chromosome 3, nga32, nga162, Arlim, GAPA, GL1, NIT1, AFC1, nga112; chromosome 4, GA1, DET1, COP9B, AG, nga1139, nga1107; and chromosome 5, CTR, ca72, nga139, SO262, SO191, DFR, ASB2, LFY3) and their map locations are described in Copenhaver *et al.* (1998). More complete information is available at the Arabidopsis Information Resource (TAIR) web site (http://www.arabidopsis.org).

Markers were scored and recorded independently by two people and then cross-checked. The complete absence of "gene conversions" in these tetrad data sets further testifies to the reliability of the scoring. The data for each of the five Arabidopsis chromosomes (Tables 1–5) were analyzed separately assuming no chromatid interference. For $(n + 1)$ markers, our data consist of tetrad patterns $(t_1, t_2, \ldots, t_n)$, where $t_i = 0$ denotes parental ditype, $t_i = 1$ denotes parental tetratype, and $t_i = 2$ denotes nonparental ditype with respect to the ith and $(i + 1)^{st}$ markers. We extend the model of Zhao *et al.* (1995) to include two types of crossover resolutions of the double-strand-break intermediates: type I without interference and type II with interference. Our model has two parameters: the interference parameter, m, which is the number of obligate "failures" between crossover "successes", and the probability, p, which is the proportion of type I (without interference) crossovers out of all crossovers. For our analyses, the intermarker genetic distances were determined by the Perkins formula; that is, $X = (\mathrm{TT}/2 + 3\mathrm{NPD})/N$, where TT is the number of tetratypes and NPD is the number of nonparental ditypes observed out of a sample of size N.

We determine the maximum-likelihood estimators for m and p from the log-likelihood function: $L(m, p) = \Sigma \log(\Pr((t_1, t_2, \ldots, t_n)|m, p, y_1, y_2, \ldots, y_n))$, where the sum is taken over all the tetrads in the data set under consideration. See the appendix for the calculation of $\Pr((t_1, t_2, \ldots, t_n)|m, p, y_1, y_2, \ldots, y_n)$.

We restrict the possibilities for the interference parameter, m, to be integers between 0 and 20 and we allow p, the probability that a randomly chosen crossover is of the noninterference type, to range between 0 and 1. For each fixed m, we determine the value of p, p_m, which maximizes the log-likelihood function, using the golden section algorithm. We then find the pair, (m, p_m), which maximizes the log-likelihood function over all the possibilities for m.

To determine whether the model with the additional parameter, p, provides a substantially better fit to the tetrad data from Arabidopsis than an interference-alone model (for which $p = 0$), we conducted a likelihood-ratio test. The test statistic is two times the difference between the maximum of the log-likelihood function under the extended model and the maximum of the log-likelihood function under the null or interference-only model. For large sample sizes, this test statistic will have approximately a chi-square distribution with degrees of freedom equal to the difference in the number of parameters involved in the extended and null models. In this case, there is one extra parameter, p, in the extended model. We verified that our data set consisting of 57 three- or four-viable spore tetrads was large enough for the distribution of the test statistic to

TABLE 1

Tetrad data for Arabidopsis chromosome 1

	Crossover pattern observed											
Intermarker distances (M):	0.272	0.105	0.061	0.114	0.132	0.053	0.228	0.149	0.114	0.088	0.044	Frequency
	0	0	0	0	0	0	1	0	0	0	0	1
	0	0	0	0	1	0	0	0	0	0	0	1
	0	0	0	1	0	0	0	0	0	0	0	2
	0	0	0	0	0	1	0	0	0	0	1	1
	0	0	0	0	1	0	0	0	1	0	0	2
	0	0	0	1	0	0	0	0	1	0	0	1
	0	0	0	1	0	0	0	1	0	0	0	1
	0	0	0	1	0	0	1	0	0	0	0	1
	0	0	1	0	0	0	0	0	1	0	0	2
	0	0	1	0	0	0	0	1	0	0	0	1
	0	1	0	0	0	0	0	0	0	1	0	1
	0	1	0	0	0	0	0	0	1	0	0	1
	0	1	0	0	0	0	0	1	0	0	0	3
	0	1	0	0	0	0	1	0	0	0	0	2
	1	0	0	0	0	0	0	1	0	0	0	2
	1	0	0	0	0	0	1	0	0	0	0	6
	1	0	0	0	0	1	0	0	0	0	0	1
	1	0	0	1	0	0	0	0	0	0	0	1
	0	0	0	0	1	0	1	1	0	0	0	1
	0	0	0	1	0	0	1	0	0	1	0	1
	0	0	1	0	1	0	0	0	0	1	0	1
	0	1	0	0	0	0	0	1	0	1	0	1
	0	1	0	0	0	0	1	1	0	0	0	1
	1	0	0	0	0	0	1	0	0	1	0	1
	1	0	0	0	0	0	1	0	1	0	0	3
	1	0	0	0	0	1	1	0	0	0	0	1
	1	0	0	0	1	0	0	0	1	0	0	1
	1	0	0	0	1	0	0	1	0	0	0	3
	1	0	0	1	0	0	0	0	0	1	0	2
	1	0	0	1	0	0	1	0	0	0	0	2
	1	0	1	0	0	0	0	0	0	0	1	1
	1	0	0	0	0	0	1	1	0	1	0	1
	1	0	0	0	0	1	1	1	0	0	0	1
	1	1	0	0	0	1	0	0	1	0	0	1
	0	1	0	0	0	0	1	1	1	1	0	1
	1	0	0	1	2	0	0	0	0	0	1	1
	1	0	1	0	0	0	1	1	0	0	1	1
	1	1	0	1	0	0	1	0	0	1	0	1
	1	0	1	0	0	1	1	0	1	0	1	1

be well approximated by a chi-square distribution with 1 d.f. by simulating data under the null hypothesis, forming the test statistic, and checking that the chi-square cut-off for rejection at the 5% significance level, $\chi^2(0.95) = 3.84$, led to rejection of the null hypothesis no more than 5% of the time.

The results of our analysis of the five linkage groups in Arabidopsis are summarized in Table 6. The model with two crossover pathways (one with and one without interference) fits the data on the longer linkage groups, 1, 3, and 5, substantially better than does the model with only an interference pathway. There is no reason to believe that the true values of the interference parameter, m, differ for these linkage groups. The distribution of the estimator m for these data sets is dispersed and skewed to the right. Due to the skewness and to computational problems encountered in obtaining estimates of $m > 20$, we cannot report meaningful standard errors or confidence intervals for the parameter estimates. However, simulations indicate that if the true interference parameter were 10, obtaining estimates for m of 17 is likely. Similarly, if the true interference parameter were 17, obtaining estimates for m of 10 is likely. On the other hand, these simulations reveal that if the true value of m were 3, estimates of 10 and 17 are unlikely and if the true value of m were 5, estimates of 10 are possible but estimates of 17 are unlikely.

The estimate of the proportion of crossovers without

TABLE 2

Tetrad data for Arabidopsis chromosome 2

	Crossover pattern observed					
Intermarker distances (M):	0.079	0.070	0.228	0.070	0.132	Frequency
	0	0	0	0	0	5
	0	0	0	1	0	6
	0	0	0	0	1	3
	0	1	0	0	0	6
	1	0	0	0	0	4
	1	0	0	1	0	2
	0	0	1	0	0	19
	0	1	0	0	1	2
	1	0	0	0	1	3
	0	0	1	0	1	7

interference, p, was bounded above by ~0.25 and generally was close to 0.20. While the estimate did occasionally fall below 0.10 when two crossover pathways were simulated, the case for $p > 0$ comes strongly from the fact that, when only the interference pathway was simulated, statistically significant estimates for $p > 0$ were rare (<5%).

Because our markers span the centromere on each chromosome, we considered the possibility that centromere disruption of interference might be the cause of our positive estimates for type I (without interference) crossovers. To rule out this possibility, we simulated data for chromosome 1 under an interference-only model (with $m = 3$ and with $m = 10$) but with complete disrup-

TABLE 3

Tetrad data for Arabidopsis chromosome 3

	Crossover pattern observed							
Intermarker distances (M):	0.149	0.228	0.132	0.061	0.167	0.219	0.175	Frequency
	0	0	0	0	0	1	0	2
	0	0	0	0	1	0	0	2
	0	0	0	1	0	0	0	2
	0	0	1	0	0	0	0	3
	0	0	0	0	1	0	1	1
	0	0	0	0	1	1	0	3
	0	0	0	1	0	0	1	1
	0	0	0	1	0	1	0	1
	0	0	1	0	0	0	1	3
	0	0	1	0	0	1	0	1
	0	1	0	0	0	0	1	6
	0	1	0	0	0	1	0	8
	0	1	0	0	1	0	0	2
	0	1	0	1	0	0	0	1
	1	0	0	0	0	0	1	2
	1	0	0	0	0	1	0	4
	1	0	0	0	1	0	0	1
	1	0	0	1	0	0	0	1
	1	0	1	0	0	0	0	1
	0	1	1	0	0	1	0	2
	0	2	0	0	0	1	0	1
	1	0	0	0	1	0	1	1
	1	0	0	0	1	1	0	1
	1	0	0	1	0	0	1	1
	1	0	1	0	0	0	1	3
	0	0	0	0	2	1	1	1
	1	1	1	0	1	0	0	1
	1	0	1	0	1	1	1	1

TABLE 4

Tetrad data for Arabidopsis chromosome 4

	Crossover pattern observed					
Intermarker distances (M):	0.132	0.149	0.070	0.193	0.044	Frequency
	0	0	0	0	0	2
	0	0	0	0	1	1
	0	0	0	1	0	14
	0	0	1	0	0	5
	0	1	0	0	0	15
	1	0	0	0	0	8
	0	0	0	1	1	1
	0	0	1	1	0	2
	0	1	0	0	1	1
	0	1	0	1	0	1
	1	0	0	0	1	2
	1	0	0	1	0	4
	1	0	1	0	0	1

tion of interference by the centromere. The null hypothesis (that the interference-only model explains the data as well as the extended model) was not rejected more often than expected by chance (5%); centromere disruption did lead to a decreased estimate for the interference parameter, *m*, on average. Thus, we conclude that centromere disruption does not explain our significant test results.

To verify that our results were not the spurious consequence of having added a parameter (*p*) to the interference-alone model, we ran the test against the Drosophila data of Bridges and Curry (Morgan *et al.* 1935) and the Neurospora data of Perkins (1962). For Neurospora, the interference-alone null model provides a good statistical fit to the data (Zhao *et al.* 1995) as well as a good visual fit (Foss *et al.* 1993). For Drosophila,

TABLE 5

Tetrad data for Arabidopsis chromosome 5

	Crossover pattern observed							
Intermarker distances (M):	0.228	0.079	0.096	0.070	0.096	0.368	0.035	Frequency
	0	0	0	0	0	1	0	6
	0	0	0	0	1	0	0	4
	0	0	0	1	0	0	0	3
	0	0	1	0	0	0	0	3
	0	1	0	0	0	0	0	1
	1	0	0	0	0	0	0	2
	0	0	0	1	0	1	0	2
	0	0	1	0	0	0	1	1
	0	0	1	0	0	1	0	3
	0	0	1	0	1	0	0	1
	0	1	0	0	0	0	1	1
	0	1	0	0	0	1	0	5
	0	1	1	0	0	0	0	1
	1	0	0	0	0	0	1	2
	1	0	0	0	0	1	0	11
	1	0	0	0	1	0	0	3
	1	0	0	1	0	0	0	1
	1	0	1	0	0	0	0	1
	1	0	0	0	0	2	0	1
	1	0	0	0	1	1	0	1
	1	0	0	1	0	1	0	1
	1	1	0	0	0	1	0	1
	1	0	0	0	1	2	0	1
	1	0	1	1	1	0	0	1

TABLE 6

Estimates of *m* and *p*

Chromosome	Null model *m* estimate	Extended model		Likelihood-ratio test statistic	*P* value[a]
		m estimate	*p* estimate		
1	3	10	0.19	14.3	0.0002
2	9	9	0.00	0	1
3	3	14	0.20	17.1	<0.0001
4	7	20	0.09	4.84	0.0288
5	3	17	0.20	15.3	0.0001

[a] Probability that the difference between the observed *p* estimate and zero could be due to sampling error alone.

the statistical fit is somewhat lacking (Zhao *et al.* 1995) although the visual fit of the interference-alone model is quite good (Foss *et al.* 1993). Tests of our extended model against the Neurospora and Drosophila data sets confirm those observations. For the Neurospora data set, the most likely value for p in the extended model is 0 and the most likely value for the interference parameter, m, is 2, in keeping with previous analyses. [The finding of $p = 0$ predicts that recombination functions are not required for synapsis in Neurospora and that Neurospora may lack a *DMC1* homolog. *mei-3*, the only *recA* homolog reported for Neurospora, belongs to the *RAD51* subfamily (Heyer 1994; Hatakeyama *et al.* 1995).] For the Drosophila data set of Bridges and Curry (Morgan *et al.* 1935), the most likely value for p in the extended model is 0.01 and the most likely value for the interference parameter, m, is 4. The visual difference between the null model and the extended model with $p = 0.01$ is virtually undetectable. The positive value for the probability that a crossover is of type I (without interference), while mildly significant (test statistic of 4.74 and a *P* value of 0.0295), is of little practical significance.

DISCUSSION

Limitations of the conclusions: Although our analysis yields results that are compatible with two discrete classes of crossovers in Arabidopsis, those with and those without interference, by themselves they are not strong support for that view. For instance, some models in which interference is imposed by a "careless" counting mechanism acting upon a single class of crossovers may not be ruled out by the data (*e.g.*, Lange *et al.* 1997). More direct evidence for two discrete pathways will likely require the isolation of mutants that specifically eliminate one or the other class, as appears to have been done for yeast.

Our analyses of chromosomes 1, 3, and 5 gave comparable estimates of m and p, with $p = 0$ ruled out. For the short chromosomes, 2 and 4, there were insufficient data to rule out a p value of 0. While the short chromosomes may not differ from the others with respect to p, it remains possible that crossing over on chromosomes 2 and 4 occurs only, or primarily, by the interference pathway. Briscoe and Tomkiel (2000) and McKee *et al.* (2000), for Drosophila, and Stitou *et al.* (2001), for rodents, have concluded that nucleolus organizers (NORs) can act as chromosome pairing centers. We noted in the Introduction the view that creatures whose chromosomes are well endowed with *cis*-acting pairing centers, like those of Drosophila and *C. elegans*, have no need for recombinational interactions to achieve synapsis. If further data demonstrate small p values for chromosomes 2 and 4, we would propose that the presence of NORs on those two chromosomes results in a reduced requirement for the noninterfering crossovers of the pairing pathway.

In Arabidopsis, Moran *et al.* (2001) noted that two different mutant strains defective in synapsis suffered reductions in chiasmata differentially on the long (1, 3, and 5) and short (2 and 4) chromosomes. It is notable that a chromosome-specific response to meiotic mutations has not been seen in Drosophila (S. Hawley, personal communication). This difference may reflect the postulated presence of two crossover pathways in Arabidopsis and one in Drosophila.

Are the large estimates for *m* realistic? Stack and Anderson (1986) reported a 15-fold excess of early recombination nodules over late nodules in tomatoes, suggestive of a large value for m. In *Zea mays*, Franklin *et al.* (1999) noted a 10- to 20-fold excess of Rad51p zygotene foci over the estimated total number of crossovers per nucleus. However, although Rad51p does promote repair of double-strand breaks by binding to ssDNA at the resected ends, these authors were disinclined to view the excess as signaling double-strand-break-induced noncrossovers: (1) "It would seem both unnecessary and catastrophic . . ."; and (2) "Copenhaver *et al.* (1998) did not observe any gene conversion events in Arabidopsis, despite scoring >1000 polymorphic loci." However, our estimate of $m \approx 15$ for Arabidopsis implies, within the framework of the counting model, that green plants may indeed have a large excess

of conversions over crossovers. We therefore calculate how many conversions should have been seen in the tetrad data involving 52 markers reported by Copenhaver *et al.* (1998). To make the calculation, we make several assumptions:

1. For each noninterference crossover there is one noncrossover in that pathway. This is equivalent to assuming that the canonical double Holliday junction intermediate is equally likely to be resolved to give a crossover or a noncrossover.
2. For each crossover in the interference pathway, the total number of potential conversions (crossovers plus noncrossovers) = $(m + 1)$. Combining assumptions (1) and (2), the ratio of total conversions to crossovers is $2p + (m + 1)(1 - p)$.
3. We take the length of a conversion tract to be 1 kb, which is about what it is in better-characterized organisms.

The probability that a gene conversion tract of 1 kb will coincide with a particular marker on a chromosome of length L is 1 kb/L. We calculated the length of each Arabidopsis chromosome by adding the number of sequenced nucleotides between the distal-most markers scored on each chromosome (Arabidopsis Genome Initiative 2000; GenBank accession nos. NC 003070, NC 003071, NC 003074, NC 003075, and NC 003076) to the sizes of the respective unsequenced centromere arrays that were estimated by fluorescence *in situ* hybridization methods (Haupt *et al.* 2001). The expected number of gene conversion tracts per bivalent (S) is 2[chromosome length in morgans][$2p + (m + 1)(1 - p)$]. Combining these two calculations, the mean number of times a gene conversion tract is expected to coincide with a given marker is (S)(1 kb/L). To derive the number of expected gene conversions for each chromosome, this value is multiplied by the number of markers scored on a given chromosome and by the number of tetrads in which a conversion could have been seen. The latter number (33) is equal to the number of full (four viable spores) tetrads (25) plus one-quarter of the number of tetrads with three viable spores (32). The resulting value was halved to account for mismatch repair processes that restore potential conversions to the Mendelian ratio, yielding 0.21, 0.08, 0.15, 0.24, and 0.21 expected gene conversions on chromosomes 1–5, respectively, for a total of 0.89 gene conversions (0.88 if the individual chromosome values are not rounded prior to totaling). If we omit the unsequenced centromere arrays from the calculation of L, we expect 0.96 gene conversions in the full data set reported in Copenhaver *et al.* (1998). We conclude that the failure to have seen a conversion in these Arabidopsis tetrad data is compatible with our estimates of the number of double-strand breaks per crossover. This conclusion reduces the force of the speculation that Rad51p is involved in early pairing exercises that are not associated with breaks.

We thank Daphne Preuss for critical discussions. Terry Speed, Scott Hawley, Edward Van Veen, and Wolf Heyer made useful suggestions. Jette Foss and Lisa Young helped groom the text. G.P.C. was supported by National Science Foundation grant DBI-9872641. E.H. was supported by an Interdisciplinary Grant in the Mathematical Sciences, DMS-0075143. Support to F.W.S. was from National Science Foundation grant MCB-9402695 and National Institutes of Health grant GM-33677.

LITERATURE CITED

Albertson, D. G., A. M. Rose and A. M. Villeneuve, 1997 Chromosome organization, mitosis, and meiosis, pp. 47–78 in *C. elegans II*, edited by D. L. Riddle, T. Blumenthal, B. J. Meyer and J. P. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Arabidopsis Genome Initiative, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815.

Bishop, D. K., D. Park, L. Xu and N. Kleckner, 1992 *DMC1*: a meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. Cell **69:** 439–456.

Borts, R. H., and J. E. Haber, 1989 Length and distribution of meiotic gene conversion tracts and crossovers in *Saccharomyces cerevisiae*. Genetics **123:** 69–80.

Borts, R. H., W.-Y. Leung, K. Kramer, B. Kramer, M. S. Williamson *et al.*, 1990 Mismatch repair-induced meiotic recombination requires the *pms1* gene product. Genetics **124:** 573–584.

Briscoe, A. Jr., and J. E. Tomkiel, 2000 Chromosomal position effects reveal different *cis*-acting requirements for rDNA transcription and sex chromosome pairing in *Drosophila melanogaster*. Genetics **155:** 1195–1211.

Celerin, M., S. T. Merino, J. E. Stone, A. M. Menzie and M. E. Zolan, 2000 Multiple roles of Spo11 in meiotic chromosome behavior. EMBO J. **19:** 2739–2750.

Chen, W., and S. Jinks-Robertson, 1999 The role of mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. Genetics **151:** 1299–1313.

Cobbs, G., 1978 Renewal process approach to the theory of genetic linkage. Case of no chromatid interference. Genetics **89:** 563–581.

Copenhaver, G. P., W. E. Brown and D. Preuss, 1998 Assaying genome-wide recombination and centromere functions with *Arabidopsis* tetrads. Proc. Natl. Acad. Sci. USA **95:** 247–252.

Couteau, F., F. Belzile, C. Horlow, O. Grandjean, D. Vezon and M. P. Doutriaux, 1999 Random chromosome segregation without meiotic arrest in both male and female meiocytes of a *dmc1* mutant of *Arabidopsis*. Plant Cell **11:** 1623–1634.

Doutriaux, M. P., F. Couteau, C. Bergounioux and C. White, 1998 Isolation and characterisation of the RAD51 and DMC1 homologs from *Arabidopsis thaliana*. Mol. Gen. Genet. **257:** 283–291.

Foss, E. J., and F. W. Stahl, 1995 A test of a counting model for chiasma interference. Genetics **139:** 1201–1209.

Foss, E., R. Lande, F. W. Stahl and C. M. Steinberg, 1993 Chiasma interference as a function of genetic distance. Genetics **133:** 681–691.

Franklin, A. E., J. McElver, I. Sunjevaric, R. Rothstein, B. Bowen *et al.*, 1999 Three-dimensional microscopy of the Rad51 recombination protein during meiotic prophase. Plant Cell **11:** 809–824.

Grelon, M., D. Vezon, G. Gendrot and G. Pelletier, 2001 AtSPO11-1 is necessary for efficient meiotic recombination in plants. EMBO J. **20:** 589–600.

Hatakeyama, S., C. Ishii and H. Inoue, 1995 Identification and expression of the *Neurospora crassa mei-3* gene which encodes a protein homologous to Rad51 of *Saccharomyces cerevisiae*. Mol. Gen. Genet. **249:** 439–446.

Haupt, W., T. C. Fischer, S. Winderl, P. Fransz and R. A. Torres-Ruiz, 2001 The Centromere1 (CEN1) region of *Arabidopsis thaliana*: architecture and functional impact of chromatin. Plant J. **27:** 285–296.

Hawley, R. S., 1980 Chromosomal sites necessary for normal levels

of meiotic recombination in *Drosophila melanogaster*. I. Evidence for and mapping of the sites. Genetics **94:** 625–646.

Heyer, W. D., 1994 The search for the right partner: homologous pairing and DNA strand exchange proteins in eukaryotes. Experientia **50:** 223–233.

Hilliker, A. J., and A. Chovnick, 1981 Further observation on intragenic recombination in *Drosophila melanogaster*. Genet. Res. **38:** 281–296.

Hilliker, A. J., S. H. Clark and A. Chovnick, 1991 The effect of DNA sequence polymorphism on intragenic recombination in the rosy locus of *Drosophila melanogaster*. Genetics **129:** 779–781.

Kleckner, N., 1996 Meiosis: how could it work? Proc. Natl. Acad. Sci. USA **93:** 8167–8174.

Klimyuk, V. I., and J. D. Jones, 1997 AtDMC1, the *Arabidopsis* homologue of the yeast *DMC1* gene: characterization, transposon-induced allelic variation and meiosis-associated expression. Plant J. **11:** 1–14.

Lande, R., and F. W. Stahl, 1993 Chiasma interference and the distribution of exchanges in *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. **58:** 543–552.

Lange, K., H. Zhao and T. P. Speed, 1997 The poisson-skip model of crossing over. Ann. Appl. Prob. **7:** 299–313.

McKee, B. D., C.-S. Hong and S. Das, 2000 On the roles of heterochromatin and euchromatin in meiosis in Drosophila: mapping chromosomal pairing sites and testing candidate mutations for effects on X-Y nondisjunction and meiotic drive in male meiosis. Genetica **109:** 77–93.

McPeek, M. S., and T. P. Speed, 1995 Modeling interference in genetic recombination. Genetics **139:** 1031–1044.

Moran, E. S., S. J. Armstrong, J. L. Santos, F. C. H. Franklin and G. H. Jones, 2001 Chiasma formation in *Arabidopsis thaliana* accession Wassileskija and in two meiotic mutants. Chromosome Res. **9:** 121–128.

Morgan, T. H., C. B. Bridges and J. Schultz, 1935 (Experiment of C. B. Bridges and V. Curry.) Carnegie Inst. Year Book **34:** 287.

Mortimer, R. K., and S. Fogel, 1974 Genetical interference and gene conversion, pp. 263–275 in *Mechanisms in Recombination*, edited by R. F. Grell. Plenum, New York.

Perkins, D. D., 1962 Crossing-over and interference in a multiply-marked chromosome arm of Neurospora. Genetics **47:** 1253–1274.

Perkins, D. D., R. Lande and F. W. Stahl, 1993 Estimates of the proportion of recombination intermediates that are resolved with crossing over in *Neurospora crassa*. Appendix to Foss *et al.* (1993). Genetics **133:** 690–691.

Preuss, D., S. Y. Rhee and R. W. Davis, 1994 Tetrad analysis possible in *Arabidopsis* with mutation of the QUARTET (QRT) genes. Science **264:** 1458–1460.

Roeder, G. S., 1997 Meiotic chromosomes: it takes two to tango. Genes Dev. **11:** 2600–2621.

Romanienko, P. J., and R. D. Camerini-Otero, 2000 The mouse *spo11* gene is required for meiotic chromosome synapsis. Mol. Cell **6:** 975–987.

Shinohara, A., H. Ogawa and T. Ogawa, 1992 Rad51 protein involved in repair and recombination in *Saccharomyces cerevisiae* is a RecA-like protein. Cell **69:** 457–470.

Stack, S. M., and L. K. Anderson, 1986 Two-dimensional spreads of synaptonemal complexes from solanaceous plants. II. Synapsis in *Lycopersicon esculentum* (tomato). Am. J. Bot. **73:** 264–281.

Stam, P., 1979 Interference in genetic crossing over and chromosome mapping. Genetics **92:** 573–594.

Stitou, S., R. Jimenez, R. Diaz De La Guardia and M. Burgos, 2001 Silent ribosomal cistrons are located at the pairing segment of the postreductional sex chromosomes of *Apodemus sylvaticus* (Rodentia, Muridae). Heredity **86:** 128–133.

Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl, 1983 The double-strand-break repair model for recombination. Cell **33:** 25–35.

Villeneuve, A. M., 1994 A *cis*-acting locus that promotes crossing over between X chromosomes in *Caenorhabditis elegans*. Genetics **136:** 887–902.

Zalevsky, J., A. J. MacQueen, J. B. Duffy, K. J. Kemphues and A. M. Villeneuve, 1999 Crossing over during *Caenorhabditis elegans* meiosis requires a conserved MutS-based pathway that is partially dispensable in budding yeast. Genetics **153:** 1271–1291.

Zhao, H., T. P. Speed and M. S. McPeek, 1995 Statistical analysis of crossover interference using the chi-square model. Genetics **139:** 1045–1056.

Communicating editor: M. E. Zolan

APPENDIX

Theorem 1. *Let X be the genetic distance spanned by an interval in morgans. Define* $\mathbf{D}(k, m, p, y)$ *to be the* $(m + 1) \times (m + 1)$ *matrix with* (i, j) *entry given by*

$$\sum_{l=0}^{k} \frac{e^{-y} y^n}{n!} \binom{n}{l} \left(\frac{p}{p + (m+1)(1-p)}\right)^{l} \left(\frac{(m+1)(1-p)}{p + (m+1)(1-p)}\right)^{n-l} \delta_{(l<k \text{ or } j \geq i)},$$

where $n = (j - i) + k + m(k - l)$, $y = 2(p + (m + 1)(1 - p))X$ *is the rate of the Poisson events (type I crossovers* C_x'*, type II crossovers* C_x''*, and type II simple gene conversions* C_o''*), and* $\delta_{(l<k \text{ or } j \geq i)}$ *is* 1 *if* $l < k$ *or* $j \geq i$ *and* 0 *otherwise. Define*

$$\mathbf{P}(m, p, y) = \mathbf{D}(0, m, p, y) + \sum_{s=2}^{\infty} (1/3)(1/2 + (-1/2)^s)\mathbf{D}(s, m, p, y)$$

$$\mathbf{T}(m, p, y) = \mathbf{D}(1, m, p, y) + \sum_{s=2}^{\infty} (2/3)(1 - (-1/2)^s)\mathbf{D}(s, m, p, y)$$

and

$$\mathbf{N}(m, p, y) = \sum_{s=2}^{\infty} (1/3)(1/2 + (-1/2)^s)\mathbf{D}(s, m, p, y).$$

Then the probability of tetrad pattern $(t_1, t_2, \ldots, t_n)$ *given the transformed intermarker distances* $y_1\, y_2, \ldots, y_n$*, interference parameter,* m*, and probability a crossover is of type I,* p*, is given by*

$$\Pr((t_1, t_2, \ldots, t_n) | y_1, y_2, \ldots, y_n, m, p) = \frac{1}{m+1} \mathbf{1}\mathbf{M}_1\mathbf{M}_2 \cdots \mathbf{M}_n 1',$$

where $\mathbf{M}_v = \mathbf{P}(m, p, y_v)$ *if* $t_v = 0$ *(parental ditype),* $\mathbf{M}_v = \mathbf{T}(m, p, y_v)$ *if* $t_v = 1$ *(tetratype), and* $\mathbf{M}_v = \mathbf{N}(m, p, y_v)$ *if* $t_v = 2$ *(nonparental ditype).*

Proof. Let m be the interference parameter, *i.e.*, the number of type II simple gene conversions between two type II crossovers. Let p be the probability any particular crossover is a type I crossover and $1 - p$ be the probability any particular crossover is a type II crossover.

Let y be the standardized interval length (standardized so that the rate for all Poisson events is 1). Since the rate for a tetrad is twice that for a bivalent, type I crossovers make up a fraction p of the crossovers in the interval and are an independent portion of the Poisson events, type II crossovers make up a fraction $(1 - p)$ of the interval and for every type II crossover we expect m type II simple gene conversions [that is, we see only a fraction, $1/(m + 1)$, of the type II Poisson events], the rate for all Poisson events is $2(p + (1 - p)(m + 1))$ and $y = 2(p + (1 - p)(m + 1))X$.

We want to form a $(m + 1) \times (m + 1)$ matrix, $\mathbf{D}(k, m, p, y)$, whose (i, j) entry $d_{i,j}(k, m, p, y)$ is the probability of having k crossovers in the current interval of length y and j type II simple gene conversions C_o''*'s* after the last type II crossover C_x'' in the current interval given

that we have i type II simple gene conversions C_o''''s after the last type II crossover C_x'' in the previous interval. Note that the k crossovers have to be distributed between type I (C_x') and type II (C_x'') crossovers.

Let l ($0 \le l \le k$) be the number of type I crossovers so that $k - l$ of the crossovers are of type II. To count the number of Poisson events, n, that we will have in the current interval, note that we need $m - i$ type II simple gene conversions (C_o''''s) before we get the first type II crossover (C_x''), the l type I crossovers (C_x''s), the first type II crossover (C_x''), $k - l - 1$ patterns of m C_o''''s followed by a C_x'', and then j C_o''''s. Thus, $n = (m - i) + l + 1 + (m + 1)(k - l - 1) + j = (j - i) + k + m(k - l)$.

Also note that Pr(n Poisson events *and* l type I events in the current interval) = Pr(l type I events in the current interval *given* n Poisson events)Pr(n Poisson events). The distribution of type I crossovers given n Poisson events is just the binomial distribution. The probability that any given Poisson event is a type I event is the ratio of the rate of type I's to the rate of all events: $p/(p + (1 - p)(m + 1))$. Thus,

Pr(l type I events in the current interval *given* n Poisson events)

$$= \binom{n}{l}\left(\frac{p}{p + (m + 1)(1 - p)}\right)^{l}\left(\frac{(m + 1)(1 - p)}{p + (m + 1)(1 - p)}\right)^{n-l}.$$

The probability of having n Poisson events in the current interval is just

$$\frac{e^{-y}y^{n}}{n!}.$$

In the special case where $l = k$, all crossovers in the interval are of type I and none are of type II. In this case, we have that j, the number of simple type II gene conversions after the last type I crossover in the current interval, must be at least i, the number of simple type II gene conversions after the last type I crossover in the previous interval, since we had no type II crossovers in the current interval.

Thus the general formula for the (i, j) entry in $\mathbf{D}(k, m, p, y)$ is

$$d_{ij}(k) = \sum_{l=0}^{k}\frac{e^{-y}y^{n}}{n!}\binom{n}{l}\left(\frac{p}{p + (m + 1)(1 - p)}\right)^{l} \times \left(\frac{(m + 1)(1 - p)}{p + (m + 1)(1 - p)}\right)^{n-l}\delta_{(l<k \text{ or } j\ge i)},$$

where $n = (j - i) + k + m(k - l)$.

The sum of all these probabilities over all the possibilities for the number of type I crossovers gives the probability of having k crossovers and j type II simple gene conversions C_o''''s after the last type II crossover in the current interval given i type II simple gene conversions C_o''''s after the last type II crossover in the preceding interval. Thus, for instance, $1/(m + 1)(1, 1, \ldots, 1)$ $\mathbf{D}(k, m, p, y)$ $(1, 1, \ldots, 1)'$, the sum over all the possibilities in the preceding and current interval for the number of C_o''''s after the last type II crossover, with each preceding possibility equally likely, gives the probability of having exactly k crossovers of any type in an interval of transformed length y.

Similarly, $1/(m + 1)(1, 1, \ldots, 1)$ $\mathbf{D}(k_1, m, p, y_1)$ $\mathbf{D}(k_2, m, p, y_2)$ $(1, 1, \ldots, 1)'$ is the sum over all preceding and ending possibilities for the number of C_o''''s after the last type II crossover for two adjacent intervals, giving the probability of having k_1 crossovers in the first interval of transformed length y_1 and k_2 crossovers in the second, adjacent, interval of transformed length y_2.

Given $k \ge 1$ crossovers in an interval and assuming no chromatid interference, the probability the resulting tetrad pattern would be $t = 0$ (parental ditype) is equal to the probability that the pattern would be $t = 2$ (nonparental ditype) and is $(1/3)(1/2 + (-1/2)^k)$; the probability the tetrad pattern would be $t = 1$ (tetratype) is thus $(2/3)(1 - (-1/2)^k)$ (Zhao *et al.* 1995).

Thus, the (i, j) entries of the matrices $\mathbf{P}$ if $t = 0$, $\mathbf{T}$ if $t = 1$, and $\mathbf{N}$ if $t = 2$ give the probability of having the specified tetrad type and j type II simple gene conversions C_o''''s after the last type II crossover in the current interval given i type II simple gene conversions C_o''''s after the last type II crossover in the preceding interval. Thus the probability of having a specified tetrad pattern is as claimed. Q.E.D.

For the analysis, we did not determine the maximum-likelihood estimators for the genetic distances between markers. Instead, we used the formula

$$X = \frac{\mathrm{TT}/2 + 3\mathrm{NPD}}{N},$$

where TT is the total number of tetratypes in the interval, NPD is the total number of nonparental ditypes in the interval, and N is the total number of tetrads scored. We used these distance estimates as our fixed intermarker distances while finding the maximum-likelihood estimators for the interference parameter, m, and the probability a crossover is of type I, p. The SAS code for conducting these analyses is available from E.A.H.